

# The Siblings of the Coupon Collector

Aristides V. Doumas  
Department of Mathematics  
National Technical University of Athens  
Zografou Campus  
157 80 Athens, GREECE  
[aris.doumas@hotmail.com](mailto:aris.doumas@hotmail.com)

and

Vassilis G. Papanicolaou  
Department of Mathematics  
National Technical University of Athens  
Zografou Campus  
157 80 Athens, GREECE  
[papanico@math.ntua.gr](mailto:papanico@math.ntua.gr)

## Abstract

The following variant of the collector's problem has attracted considerable attention relatively recently (see, e.g., [24], [14], [15], [1], and [26], listed here in chronological order): There is one main collector who collects coupons. Assume there are  $N$  different types of coupons with, in general, unequal occurring probabilities. When the main collector gets a "double", she gives it to her older brother; when this brother gets a "double", he gives it to the next brother, and so on. Hence, when the main collector completes her collection, the album of the  $j$ -th sibling,  $j = 2, 3, \dots$ , will still have  $U_j^N$  empty spaces. In this article we develop techniques of computing asymptotics of the average  $E[U_j^N]$  of  $U_j^N$  as  $N \rightarrow \infty$ , for a large class of families of coupon probabilities. We also give various illustrative examples.

**Keywords.** Urn problems; (generalized) coupon collector's problem (GCCP); asymptotics.

**2010 AMS Mathematics Classification.** 60F05; 60F99.

# 1 Introduction

## 1.1 Preliminaries

The classical “**coupon collector’s problem**” (CCP) concerns a population (e.g. fishes, viruses, genes, words, baseball cards, etc.) whose members are of  $N$  different *types*. The members of the population are sampled independently with replacement and their types are recorded. CCP pertains to the family of urn problems along with other famous problems, such as the birthday, or occupancy. Its origin can be traced back to De Moivre’s treatise *De Mensura Sortis* of 1712 (see, e.g., [18]) and Laplace’s pioneering work *Theorie Analytique de Probabilites* of 1812 (see [8]). The problem became popular in the 1930’s when the Dixie Cup company introduced a highly successful program by which children collected Dixie lids to receive “Premiums,” beginning with illustrations of their favored Dixie Circus characters, and then Hollywood stars and major league baseball players (for the Dixie Cup company history see [28]).

For  $1 \leq k \leq N$  let  $p_k$  be the probability that a member of the population is of type  $k$ ,  $\sum_{k=1}^N p_k = 1$ . Let  $T_N$  be the number of trials it takes until all  $N$  types are detected (at least once).

General results for the simplest CCP (i.e. the case of equal probabilities) had appeared in some relatively unknown works (see [21], [16], where the entertaining term *cartophily* appeared in the title of these papers). However, some of the classical references for this case of the CCP are W. Feller’s well known work [13], D.J. Newman’s and L. Shepp’s paper on the *Double Dixie Cup* problem [23] (where they answered the question: how long, on average does it take to obtain  $m$  complete sets of  $N$  coupons), and a paper of P. Erdős and A. Rényi, where the limit distribution of the random variable  $T_N$  has been given (see [12]). Since then, CCP has attracted the attention of various researchers due to the fact that it has found many applications in many areas of science (computer science/search algorithms, mathematical programming, optimization, learning processes, engineering, ecology, as well as linguistics—see, e.g., [4], [20]).

For the general case of unequal probabilities regarding the asymptotics of the moments, as well as for the limit distribution of  $T_N$ , there is a plethora of references (see for instance [2], [7], [19], [18], [5], [11], [22], [9], [10], and [3]).

A generalized and interesting version of the classical CCP assumes (see, e.g., [24], [14], [15], [1], and [26]), that the Dixie Cup company sells ice cream with a cardboard cover that has hidden on the underside a picture (“coupon”) of a sixties music band. In total there are  $N$  different pictures and each one appears with probability  $p_k$ . Mr. and Mrs. Smith have one daughter and  $(r - 1)$  sons, all ice cream and sixties music addicts. The girl (she is the oldest) is the only one to buy ice cream. She tries to complete

her collection. When she gets a new picture she puts it in her album, and when she gets a double, she gives it to her oldest brother, and when this one gets a double, he gives it to the remaining oldest brother, and so on. After having bought  $T_N$  ice creams, the girl has completed her album while they remain  $U_j^N$  unfilled places in the album of the  $j$ -th collector  $j = 2, 3, \dots, r$  (that is the  $(j - 1)$ -th brother). Obviously,

$$1 \leq U_j^N \leq N \quad \text{and} \quad U_{j-1}^N \leq U_j^N, \quad j = 2, 3, \dots,$$

where, for completeness we have used the convention that  $U_1^N = 0$ . We shall refer to such a version of coupon collection as the **Generalized CCP (GCCP)**. In this paper we study the asymptotics of the expectation of the random variable  $U_j^N$  as  $N \rightarrow \infty$ .

## 1.2 The case of equal probabilities

Naturally, the simplest case occurs when one takes

$$p_1 = \dots = p_N = 1/N.$$

This case has been studied for quite a while. For  $j = 2$ , Pintacuda (see [24]) used the martingale stopping theorem and proved that

$$E[U_2^N] = \sum_{m=1}^N \frac{1}{m} =: H_N \quad (1.1)$$

( $H_N$  is sometimes called the  $N$ -th harmonic number).

**Remark 1.** It is well known (see, e.g., [13]) that

$$E[T_N] = NH_N. \quad (1.2)$$

Hence, for large  $N$ , when the main collector has completed her collection (notice that by (1.2) she will need in average  $N \ln N + O(N)$  trials in order to succeed), the expected number of unfilled coupons in her oldest brother collection will be  $\ln N + O(1)$ .

Foata *et al.*, and Foata and Zeilberger (see [14] and [15]) using nonelementary mathematics, obtained recursive formulae for  $E[U_j^N]$ ,  $j \geq 2$ . Soonafter, Adler *et al.* [1] derived the same recursion, as well as a closed-form expression for  $E[U_j^N]$  by using basic probability arguments (again for  $j \geq 2$ , while all  $p_k$ 's were considered equal). In particular, they proved that

$$E[U_2^N] = \sum_{m=1}^N \frac{1}{m}, \quad E[U_j^N] = \sum_{m=1}^N \frac{E[U_{j-1}^m]}{m} \quad \text{for } j \geq 3. \quad (1.3)$$

and also that

$$E[U_j^N] = \sum_{k=1}^N \binom{N}{k} \frac{(-1)^{k+1}}{k^{j-1}} \quad \text{for } j = 2, 3, \dots \quad (1.4)$$

Foata *et al.* (in [14]) called the quantities appearing in the recursion of (1.3) *hyperharmonic numbers*. Notice that, for fixed  $j$ , detailed asymptotics for hyperharmonic numbers can be derived (e.g., via the associated generating functions). For example,

$$E[U_j^N] \sim \frac{(\ln N)^{j-1}}{(j-1)!}. \quad (1.5)$$

Furthermore, in the case of equal coupon probabilities, by exploiting the techniques of [1] one can compute explicitly the variance  $V[U_j^N]$  and its asymptotics as  $N \rightarrow \infty$ . In particular, for  $j = 2$  we get

$$V[U_2^N] = 4 \left( H_1 + \frac{H_2}{2} + \dots + \frac{H_N}{N} \right) - 3H_N - H_N^2, \quad (1.6)$$

(a slightly different, albeit equivalent, form of formula (1.6) can be found in [25] and [26]) hence

$$V[U_2^N] \sim \ln^2 N \quad \text{as } N \rightarrow \infty.$$

### 1.3 The case of unequal probabilities

Let us now suppose that each coupon appears with probability  $p_k$ , with  $\sum_{k=1}^N p_k = 1$ ,  $p_k > 0$  for  $k = 1, \dots, N$ . Then, Adler *et al.* [1], (see also Ross [26]) proved that, when the main collector has a complete set, the expected number of unfilled coupons in each of her brothers' albums is obtained from

$$E[U_j^N] = \sum_{k=1}^N \int_0^\infty p_k e^{-p_k t} \frac{(p_k t)^{j-1}}{(j-1)!} \left[ \prod_{i \neq k} (1 - e^{-p_i t}) \right] dt, \quad j \geq 2. \quad (1.7)$$

For example,

$$E[U_j^2] = 2 - p_1^j - p_2^j, \quad j \geq 2, \quad (1.8)$$

and

$$E[U_2^3] = 3 + p_1^2 + p_2^2 + p_3^2 - \frac{p_1^2 + p_2^2}{(p_1 + p_2)^2} - \frac{p_1^2 + p_3^2}{(p_1 + p_3)^2} - \frac{p_2^2 + p_3^2}{(p_2 + p_3)^2}. \quad (1.9)$$

In addition, in [1], the authors derived two sets of general bounds for  $E[U_j^N]$ , as well as a simulation approach for estimating the summands of (1.7).

**Conjecture.** For a fixed  $N$ , the average  $E[U_j^N]$  of  $U_j^N$  becomes maximum when all the  $p_k$ 's become equal (to  $1/N$ ).

The results of the present paper support this conjecture.

## 1.4 Large $N$ asymptotics

When  $N$  is large it is not clear at all what information one can obtain from formula (1.7). For this reason there is a need to develop efficient ways for deriving asymptotics for  $E[U_j^N]$  as  $N \rightarrow \infty$ .

Let  $\alpha = \{a_k\}_{k=1}^\infty$  be a sequence of strictly positive numbers. Then, for each integer  $N > 0$ , one can create a probability measure  $\pi_N = \{p_1, \dots, p_N\}$  on the set of types  $\{1, \dots, N\}$  by taking

$$p_k = \frac{a_k}{A_N}, \quad \text{where } A_N = \sum_{k=1}^N a_k. \quad (1.10)$$

Notice that  $p_k$  depends on  $\alpha$  and  $N$ , thus, given  $\alpha$ , it makes sense to consider the asymptotic behavior of  $E[U_j^N]$  as  $N \rightarrow \infty$ . This approach for creating sequences of probability measures was first introduced in [5] and adopted in [9] and [10].

The sequence of measures  $\pi_N$ ,  $N = 2, 3, \dots$ , constructed from  $\alpha$  via (1.10) has an interesting property:

For each  $N = 2, 3, \dots$  let

$$\Omega_N := \{\omega = \{\omega_j\}_{j=1}^\infty : \omega_j = 1 \text{ or } 2 \text{ or } \dots \text{ or } N\}, \quad (1.11)$$

i.e.  $\Omega_N = \{1, \dots, N\}^\mathbb{N}$ , where  $\mathbb{N} = \{1, 2, \dots\}$ . The sample space  $\Omega_N$  describes the experiment of sampling  $N$  coupons with replacement, indefinitely.

To each  $m$ -tuple  $(k_1, \dots, k_m) \in \{1, \dots, N\}$  ( $m \in \mathbb{N}$ ) we associate a cylinder subset of  $\Omega_N$ :

$$A_{k_1, \dots, k_m} := \{k_1\} \times \dots \times \{k_m\} \times \prod_{j=m+1}^\infty \{1, \dots, N\}. \quad (1.12)$$

Then,  $\pi_N = \{p_1, \dots, p_N\}$ , as defined in (1.10), induces a set function (a probability)  $P_N$  on these cylinder sets:

$$P_N \{A_{k_1, \dots, k_m}\} := p_{k_1}^N \cdots p_{k_m}^N \quad (1.13)$$

(here  $N$  is a superscript indicating the dependence of  $p_k$  on  $N$ ). By the Extension Theorem of Caratheodory,  $P_N$  extends to a complete probability measure (which we also denote by  $P_N$ ) on  $(\Omega_N, \mathcal{F}_N)$ , where  $\mathcal{F}_N$  is the completion of the  $\sigma$ -algebra generated by the cylinder sets.

If  $L_{N+1}$  is the subset of  $\Omega_{N+1}$  defined by

$$L_{N+1} := \{\omega = \{\omega_j\}_{j=1}^\infty : \omega_j = N+1 \text{ for all but finitely many } j\}, \quad (1.14)$$

then it is clear that

$$P_{N+1} \{L_{N+1}\} = 0. \quad (1.15)$$

For a sequence  $\omega = \{\omega_j\}_{j=1}^\infty \in \Omega_{N+1} \setminus L_{N+1}$  we denote by  $\iota_{N+1}(\omega)$  the sequence (in  $\Omega_N$ ) which is obtained from  $\omega$  by deleting all terms  $\omega_j$  such that  $\omega_j = N + 1$  (notice that  $\iota_{N+1}(\omega)$  is a sequence, i.e. it has infinitely many terms, since  $\omega \notin L_{N+1}$ ).

To be more precise, given  $\omega = \{\omega_j\}_{j=1}^\infty$  let us consider the set of indices

$$\{j_n \in \mathbb{N} : \omega_{j_n} \neq N + 1, \quad j_n < j_{n+1}\}.$$

Then,  $\iota_{N+1}(\omega) = \{\omega_{j_n}\}_{n=1}^\infty$ . We call  $\iota_{N+1}(\omega)$  the  $(N + 1)$ -contraction of  $\omega \in \Omega_{N+1} \setminus L_{N+1}$ . For example, if

$$\omega = (3, 1, 2, 3, 3, 2, 1, 2, 2, 3, 1, 1, \dots) \in \Omega_3 \setminus L_3,$$

then

$$\iota_3(\omega) = (1, 2, 2, 1, 2, 2, 1, 1, \dots) \in \Omega_2.$$

Since  $\Omega_N \subset (\Omega_{N+1} \setminus L_{N+1})$ , we have that  $\iota_{N+1} : (\Omega_{N+1} \setminus L_{N+1}) \rightarrow \Omega_N$  is a “projection” and, via  $\iota_{N+1}$ ,  $P_{N+1}$  induces a measure  $\tilde{P}_N$  on  $\mathcal{F}_N$  given by

$$\tilde{P}_N\{A\} := P_{N+1}\{\iota_{N+1}^{-1}(A)\}. \quad (1.16)$$

**Proposition 1.** The measures  $P_N$  and  $\tilde{P}_N$  coincide, namely

$$\tilde{P}_N\{A\} = P_N\{A\} \quad \text{for all } A \in \mathcal{F}_N.$$

The proof is straightforward.

**Remark 2.** For  $\omega = \{\omega_j\}_{j=1}^\infty \in \Omega_{N+1} \setminus L_{N+1}$  we have that  $T_{N+1}(\omega) =$  the smallest  $k$  such that  $(\omega_1, \dots, \omega_k)$  contains all elements of  $\{1, \dots, N + 1\}$ . If we set  $\tilde{T}_N(\omega) := T_N(\iota_{N+1}(\omega))$ , then, obviously  $\tilde{T}_N(\omega) \leq T_{N+1}(\omega)$ . Also, it is easy to see that  $\tilde{T}_N$  and  $T_N$  have the same distribution. Therefore the sequence  $\{T_N\}_{N=2}^\infty$  is *stochastically increasing*, i.e.  $P\{T_{N+1} \geq k\} \geq P\{T_N \geq k\}$  for all  $k$  and all  $N \geq 2$ .

This is not true however, for the sequence  $\{U_j^N\}_{N=2}^\infty$ . For instance, let  $\alpha = (1, 1, \epsilon, \dots)$  where  $\epsilon > 0$ . Then, for  $N = 2$  formula (1.10) gives that  $p_1 = p_2 = 1/2$  and hence, by (1.8) (or (1.1)) we get that  $E[U_2^2] = 3/2$ . On the other hand, if  $N = 3$ , then (1.10) gives  $p_1 = p_2 = 1/(2 + \epsilon)$  and  $p_3 = \epsilon/(2 + \epsilon)$ . It, then, follows from (1.9) that  $E[U_3^2]$  can be made as close to 1 as we wish, by taking  $\epsilon$  sufficiently close to 0 (as it is intuitively expected, since one coupon is very “rare” compare to the others). In particular, we can have  $E[U_3^2] < E[U_2^2]$ , for a given sequence  $\alpha$ .

If  $\alpha$  is the constant sequence whose terms are (all) equal to 1 (i.e. in the case of equal coupon probabilities), then (1.3) implies that  $E[U_j^N]$  increases with  $N$ . We conjecture that in this case we also have that, for any  $j \geq 2$ ,  $U_j^N$  is stochastically increasing with  $N$ .

After the above deviation we return to our main issue, namely the asymptotics of  $E[U_j^N]$ . Inspired by (1.7) we introduce the notation

$$I_N(\alpha; j) := \sum_{k=1}^N \int_0^\infty a_k e^{-a_k t} \frac{(a_k t)^{j-1}}{(j-1)!} \prod_{i \neq k} (1 - e^{-a_i t}) dt, \quad j \geq 2. \quad (1.17)$$

If  $s\alpha := \{sa_k\}_{k=1}^\infty$ , (1.17) gives immediately that

$$I_N(s\alpha) = I_N(\alpha) \quad (1.18)$$

and hence, in view of (1.7) and (1.10)

$$E[U_j^N] = I_N(\alpha; j). \quad (1.19)$$

The rest of the paper is organized as follows. In Section 2 we consider classes of decaying sequences  $\alpha$  such that  $a_k \rightarrow 0$ . Here the computations are quite involved. We present the main result in Theorem 1 (of Subsection 2.3). In particular, the first three terms of the asymptotic expansion of  $E[U_j^N]$  are determined (as  $N \rightarrow \infty$ ). It is notable that the *generalized Zipf law* falls in this category. The method of proving Theorem 1 is based on a “brute force” technique reminding the technique initiated in [5] and exploited in [9]. In Section 3 we derive the leading behavior of  $E[U_j^N]$  for a large class of sequences  $\alpha$ , such that  $a_k \rightarrow \infty$ . In this case we often get that  $E[U_j^N]$  approaches a finite limit as  $N \rightarrow \infty$ . Various examples are exhibited. In particular, we cover some important families of coupon probabilities (e.g. polynomial and exponential).

## 2 Decaying sequences

Inspired by [9] we consider sequences  $\alpha = \{a_k\}_{k=1}^\infty$  of the form

$$a_k = \frac{1}{f(k)}, \quad (2.1)$$

where

$$f(x) > 0 \quad \text{and} \quad f'(x) > 0, \quad x > 0, \quad (2.2)$$

and furthermore we assume that  $f(x)$  possesses three derivatives and satisfies the following conditions as  $x \rightarrow \infty$ :

$$\begin{aligned} \text{(i)} \quad & f(x) \rightarrow \infty, & \text{(ii)} \quad & \frac{f'(x)}{f(x)} = o\left(\frac{1}{\ln^p x}\right) \text{ for any } p > 0, \\ \text{(iii)} \quad & \frac{f''(x)/f'(x)}{f'(x)/f(x)} = O(1), & \text{(iv)} \quad & \frac{f'''(x) f(x)^2}{f'(x)^3} = O(1) \end{aligned} \quad (2.3)$$

(in [9] the conditions on  $f(x)$  were slightly weaker). These conditions are satisfied by a variety of commonly used functions. For example,

$$f(x) = x^p(\ln x)^q, \quad p > 0, \quad q \in \mathbb{R}, \quad f(x) = \exp(x^r), \quad 0 < r < 1,$$

or various convex combinations of products of such functions.

**Remark 3.** Condition (ii) of (2.3) implies  $f'(x)/f(x) \rightarrow 0$ . Thus,

$$\lim_{x \rightarrow \infty} \frac{f(x+1)}{f(x)} = 1 \quad (2.4)$$

(this can be justified, e.g., by applying the Mean Value Theorem to the function  $g(x) = \ln f(x)$  on the interval  $[x, x+1]$ ).

For typographical convenience we set

$$F(x) := f(x) \ln \left( \frac{f(x)}{f'(x)} \right) \quad (2.5)$$

(notice that (2.2) and (ii) of (2.3) imply that  $F(x) > 0$  for  $x$  sufficiently large). Starting from (1.17), we substitute  $t = F(N)s$  in the integral and rewrite  $I_N(\alpha; j)$  as

$$I_N(\alpha; j) = \frac{1}{(j-1)!} F(N)^j [I_N^1(\alpha; j) + I_N^2(\alpha; j)], \quad (2.6)$$

where

$$I_N^1(\alpha; j) := \sum_{k=1}^N f(k)^{-j} \int_0^1 e^{-\frac{F(N)}{f(k)}s} \exp \left[ \sum_{i \neq k} \ln \left( 1 - e^{-\frac{F(N)}{f(i)}s} \right) \right] s^{j-1} ds \quad (2.7)$$

and

$$I_N^2(\alpha; j) := \sum_{k=1}^N f(k)^{-j} \int_1^\infty e^{-\frac{F(N)}{f(k)}s} \exp \left[ \sum_{i \neq k} \ln \left( 1 - e^{-\frac{F(N)}{f(i)}s} \right) \right] s^{j-1} ds. \quad (2.8)$$

In order to analyze deeper the above quantities we need the following lemma.

**Lemma 1.** Set

$$J_m(N) := \int_1^N f(x)^m e^{-\frac{F(N)}{f(x)}s} dx, \quad m \geq 0. \quad (2.9)$$



Then, under (2.3) and (2.5), we have, as  $N \rightarrow \infty$ ,

$$J_m(N) = \frac{f(N)^{m+2}}{sF(N)f'(N)} e^{-\frac{F(N)}{f(N)}s} + \omega(N) \frac{f(N)^{m+3}}{s^2F(N)^2f'(N)} e^{-\frac{F(N)}{f(N)}s} \left[ 1 + O\left(\frac{f(N)}{F(N)}\right) \right], \quad (2.10)$$

uniformly in  $s \in [s_0, \infty)$ , for any  $s_0 > 0$ , where

$$\omega(N) := -2 + \frac{f''(N)/f'(N)}{f'(N)/f(N)}. \quad (2.11)$$

For the proof see [9]. Notice that the condition (iii) of (2.3) says that  $\omega(N) = O(1)$  as  $N \rightarrow \infty$ .

**Remark 4.** It is straightforward to check that Lemma 1 is still valid when  $m$  is a negative integer.

## 2.1 The integral $I_N^1(\alpha; j)$

Regarding the quantity of (2.7), given  $\varepsilon \in (0, 1/2)$  we have

$$I_N^1(\alpha; j) = I_N^{11}(\alpha; j) + I_N^{12}(\alpha; j), \quad (2.12)$$

where

$$I_N^{11}(\alpha; j) := \sum_{k=1}^N f(k)^{-j} \int_0^{1-\varepsilon} e^{-\frac{F(N)}{f(k)}s} \exp \left[ \sum_{i \neq k} \ln \left( 1 - e^{-\frac{F(N)}{f(i)}s} \right) \right] s^{j-1} ds \quad (2.13)$$

and

$$I_N^{12}(\alpha; j) := \sum_{k=1}^N f(k)^{-j} \int_{1-\varepsilon}^1 e^{-\frac{F(N)}{f(k)}s} \exp \left[ \sum_{i \neq k} \ln \left( 1 - e^{-\frac{F(N)}{f(i)}s} \right) \right] s^{j-1} ds. \quad (2.14)$$

The quantity of (2.13) becomes (in view of (2.2))

$$\begin{aligned} I_N^{11}(\alpha; j) &= \sum_{k=1}^N f(k)^{-j} \int_0^{1-\varepsilon} \frac{e^{-\frac{F(N)}{f(k)}s}}{1 - e^{-\frac{F(N)}{f(k)}s}} \exp \left[ \sum_{i=1}^N \ln \left( 1 - e^{-\frac{F(N)}{f(i)}s} \right) \right] s^{j-1} ds \\ &< \sum_{k=1}^N f(k)^{-j} \int_0^{1-\varepsilon} \frac{e^{-\frac{F(N)}{f(k)}s}}{1 - e^{-\frac{F(N)}{f(k)}s}} \exp \left[ \sum_{i=1}^N \ln \left( 1 - e^{-\frac{F(N)}{f(i)}(1-\varepsilon)} \right) \right] s^{j-1} ds \\ &< \exp \left( - \sum_{i=1}^N e^{-\frac{F(N)}{f(i)}(1-\varepsilon)} \right) \sum_{k=1}^N f(k)^{-j} \int_0^{1-\varepsilon} \frac{e^{-\frac{F(N)}{f(k)}s}}{1 - e^{-\frac{F(N)}{f(k)}s}} s^{j-1} ds, \end{aligned} \quad (2.15)$$

since  $\ln(1-x) < -x$ , for  $0 < x < 1$ . Now,  $f$  is increasing, hence, from the comparison of sums and integrals we have

$$\begin{aligned} \int_1^N e^{-\frac{F(N)}{f(x)}s} dx &\leq \sum_{j=1}^N e^{-\frac{F(N)}{f(j)}s} \\ &\leq \int_1^{N+1} e^{-\frac{F(N)}{f(x)}s} dx \\ &\leq \int_1^N e^{-\frac{F(N)}{f(x)}s} dx + e^{-\frac{F(N)}{f(N+1)}s}. \end{aligned} \quad (2.16)$$

Using the above comparison, (2.4), and applying Lemma 1, for  $m = 0$ , one arrives at

$$\begin{aligned} \exp\left(-\sum_{i=1}^N e^{-\frac{F(N)}{f(i)}(1-\varepsilon)}\right) &\leq \exp\left(-\int_1^N e^{-\frac{F(N)}{f(x)}(1-\varepsilon)} dx\right) \\ &< \exp\left\{-\frac{f(N)^2}{(1-\varepsilon)F(N)f'(N)} e^{-\frac{F(N)}{f(N)}(1-\varepsilon)} \left(1 + M_1 \frac{f(N)}{F(N)}\right)\right\} \\ &= \exp\left\{-\frac{1}{1-\varepsilon} \cdot \frac{\left(\frac{f(N)}{f'(N)}\right)^\varepsilon}{\ln\left(\frac{f(N)}{f'(N)}\right)} \left(1 + \frac{M_1}{\ln\left(\frac{f(N)}{f'(N)}\right)}\right)\right\}, \end{aligned} \quad (2.17)$$

where  $M_1$  is a positive constant and we have used (2.5), i.e. the definition of  $F$ . On the other hand, if we set

$$J_N(k; \varepsilon) := \int_0^{1-\varepsilon} \frac{e^{-\frac{F(N)}{f(k)}s}}{1 - e^{-\frac{F(N)}{f(k)}s}} s^{j-1} ds,$$

then the scaling  $u = F(N)f(k)^{-1}s$ , via the definition of  $F$  and the monotonicity (and positivity) of  $f$ , yields

$$\begin{aligned} J_N(k; \varepsilon) &= \left[\frac{f(k)}{F(N)}\right]^j \int_0^{(1-\varepsilon)\frac{F(N)}{f(k)}} \frac{e^{-u}}{1 - e^{-u}} u^{j-1} du \\ &< \left[\frac{f(k)}{F(N)}\right]^j \int_0^\infty \frac{e^{-u}}{1 - e^{-u}} u^{j-1} du. \end{aligned}$$

Since  $j \geq 2$ , the last integral is equal to a positive constant. In particular, for  $j = 2$ , it converges to  $\pi^2/6$ . In general, it is not difficult to check that it converges to

$$M_j := \Gamma(j) \zeta(j), \quad j = 2, 3, \dots,$$

where  $\Gamma(\cdot)$ ,  $\zeta(\cdot)$  denote the gamma function and the Riemann zeta function respectively. Hence,

$$J_N(k; \varepsilon) \leq M_j \left[\frac{f(k)}{F(N)}\right]^j. \quad (2.18)$$

In view of (2.17) and (2.18), (2.15) yields

$$I_N^{11}(\alpha; j) \leq M_j \frac{N}{F(N)^j} \exp \left\{ -\frac{1}{1-\varepsilon} \cdot \frac{\left(\frac{f(N)}{f'(N)}\right)^\varepsilon}{\ln\left(\frac{f(N)}{f'(N)}\right)} \left(1 + M_1 \frac{1}{\ln\left(\frac{f(N)}{f'(N)}\right)}\right) \right\}. \quad (2.19)$$

Since  $\varepsilon \in (0, 1/2)$ , we claim that (2.19) implies

$$I_N^{11}(\alpha; j) << \frac{\ln^2\left(\ln\left(\frac{f(N)}{f'(N)}\right)\right)}{f(N)^j \ln^3\left(\frac{f(N)}{f'(N)}\right)}, \quad j \geq 2, \quad (2.20)$$

where  $a_N << b_N$  means that  $a_N/b_N \rightarrow 0$  as  $N \rightarrow \infty$ . To check the validity of (2.20) one observes that it suffices to show that

$$N << \exp \left[ \frac{1}{1-\varepsilon} \cdot \frac{\left(\frac{f(N)}{f'(N)}\right)^\varepsilon}{\ln\left(\frac{f(N)}{f'(N)}\right)} \right] \times \left[ \frac{\ln^2\left(\ln\left(\frac{f(N)}{f'(N)}\right)\right)}{\ln^{3-j}\left(\frac{f(N)}{f'(N)}\right)} \right], \quad (2.21)$$

which follows by taking logarithms and using condition (ii) of (2.3).

Our next task is to compute a few terms of the asymptotic expansion of the term  $I_N^{12}(\alpha; j)$  defined in (2.14). For convenience we set

$$B_k(N; s) := \sum_{i=1}^N \ln \left( 1 - e^{-\frac{F(N)}{f(i)}s} \right) - \ln \left( 1 - e^{-\frac{F(N)}{f(k)}s} \right). \quad (2.22)$$

Since

$$\frac{F(N)}{f(i)} \rightarrow \infty \quad \text{as } N \rightarrow \infty,$$

and  $\ln(1-x) = -x + O(x^2)$  as  $x \rightarrow 0$ , we have (as long as  $s \geq s_0 > 0$ )

$$B_k(N; s) = \sum_{i=1}^N \left[ -e^{-\frac{F(N)}{f(i)}s} + O\left(e^{-\frac{2F(N)}{f(i)}s}\right) \right] + e^{-\frac{F(N)}{f(k)}s} + O\left(e^{-\frac{2F(N)}{f(k)}s}\right). \quad (2.23)$$

From the comparison of sums and integrals, i.e. (2.16), (2.23) yields

$$\begin{aligned} B_k(N; s) = & - \left[ \int_1^N e^{-\frac{F(N)}{f(x)}s} dx + O\left(e^{-\frac{F(N)}{f(N+1)}s}\right) \right] \\ & + \sum_{i=1}^N O\left(e^{-\frac{2F(N)}{f(i)}s}\right) + e^{-\frac{F(N)}{f(k)}s} + O\left(e^{-\frac{2F(N)}{f(k)}s}\right). \end{aligned}$$

The above formula together with Lemma 1, for  $m = 0$ , give

$$B_k(N; s) = -\frac{f(N)^2}{sF(N)f'(N)} e^{-\frac{F(N)}{f(N)}s} - \omega(N) \frac{f(N)^3}{s^2 F(N)^2 f'(N)} e^{-\frac{F(N)}{f(N)}s} \left[ 1 + O\left(\frac{f(N)}{F(N)}\right) \right]$$

$$+O\left(e^{-\frac{F(N)}{f(N+1)}s} + Ne^{-\frac{2F(N)}{f(N)}s}\right) + e^{-\frac{F(N)}{f(k)}s} + O\left(e^{-\frac{2F(N)}{f(k)}s}\right).$$

Using (2.4) the above yields

$$B_k(N; s) = -\frac{f(N)^2}{sF(N)f'(N)}e^{-\frac{F(N)}{f(N)}s} - \omega(N) \frac{f(N)^3}{s^2F(N)^2f'(N)}e^{-\frac{F(N)}{f(N)}s} \left[1 + O\left(\frac{f(N)}{F(N)}\right)\right],$$

independent of  $k$ . Hence,

$$\begin{aligned} I_N^{12}(\alpha; j) &= \sum_{k=1}^N f(k)^{-j} \int_{1-\varepsilon}^1 e^{B(N; s)} e^{-\frac{F(N)}{f(k)}s} s^{j-1} ds \\ &= \sum_{k=1}^N f(k)^{-j} \int_{1-\varepsilon}^1 e^{-\frac{F(N)}{f(k)}s} s^{j-1} \times \exp \left[ -\frac{f(N)^2}{sF(N)f'(N)} e^{-\frac{F(N)}{f(N)}s} \right. \\ &\quad \left. - \omega(N) \frac{f(N)^3}{s^2F(N)^2f'(N)} e^{-\frac{F(N)}{f(N)}s} \left[1 + O\left(\frac{f(N)}{F(N)}\right)\right] \right] ds \end{aligned}$$

as  $N \rightarrow \infty$ . Using the definition of  $F$ , namely (2.5), and substituting  $s = 1 - t$ , the above expression becomes

$$\begin{aligned} I_N^{12}(\alpha; j) &= \sum_{k=1}^N f(k)^{-j} \int_0^\varepsilon (1-t)^{j-1} \exp \left[ - (1-t) \frac{f(N)}{f(k)} \ln \left( \frac{f(N)}{f'(N)} \right) \right] \\ &\quad \times \exp \left\{ -\frac{1}{1-t} \frac{\left( \frac{f(N)}{f'(N)} \right)^t}{\ln \left( \frac{f(N)}{f'(N)} \right)} - \omega(N) \frac{1}{(1-t)^2} \frac{\left( \frac{f(N)}{f'(N)} \right)^t}{\ln \left( \frac{f(N)}{f'(N)} \right)^2} \right. \\ &\quad \left. \times \left[ 1 + O\left( \frac{1}{\ln \left( \frac{f(N)}{f'(N)} \right)} \right) \right] \right\} dt. \end{aligned}$$

For typographical convenience we set

$$A := \frac{f(N)}{f'(N)} \tag{2.24}$$

(notice that  $A \rightarrow \infty$  as  $N \rightarrow \infty$ ). Then  $I_N^{12}$  can be expressed as

$$\begin{aligned} I_N^{12}(\alpha; j) &= \int_0^\varepsilon \left[ \sum_{k=1}^N f(k)^{-j} \exp \left( -\frac{1}{f(k)} f(N) (1-t) \ln A \right) \right] \\ &\quad \times (1-t)^{j-1} \exp \left\{ -\frac{A^t}{\ln A} \left( \sum_{n=0}^\infty t^n \right) - \omega(N) \frac{A^t}{\ln^2 A} \left( \sum_{n=1}^\infty n t^{n-1} \right) \right. \\ &\quad \left. \times \left[ 1 + O\left( \frac{1}{\ln A} \right) \right] \right\} dt. \end{aligned} \tag{2.25}$$

Set

$$S_j(N; t) := \sum_{k=1}^N f(k)^{-j} e^{-\frac{f(N)}{f(k)}(1-t) \ln A}, \quad 0 \leq t \leq \varepsilon < 1/2.$$

and

$$g_j(x) := f(x)^{-j} \exp \left( -\frac{f(N)}{f(x)} (1-t) \ln A \right), \quad 0 \leq x \leq N, \quad 0 \leq t \leq \varepsilon < 1/2.$$

It is easy to check that under conditions (2.3),  $g$  is increasing for sufficiently large  $N$ . Thus, it follows from the comparison of sums and integrals that

$$S_j(N; t) = K_j(N; t) + O \left( \frac{e^{-\frac{f(N)}{f(N+1)}(1-t) \ln A}}{f(N+1)^j} \right), \quad (2.26)$$

where

$$K_j(N; t) := \int_1^N f(x)^{-j} e^{-\frac{f(N)}{f(x)}(1-t) \ln A} dx. \quad (2.27)$$

By using Lemma 1 and Remarks 4 and 5 (as long as  $1-t > 1/2$ ), we get as  $N \rightarrow \infty$ ,

$$\begin{aligned} S_j(N; t) &= \frac{A^{t-1}}{(1-t) \ln A} \cdot \frac{1}{f(N)^{j-1} f'(N)} \\ &\quad + \omega(N) \frac{A^{t-1}}{(1-t)^2 \ln^2 A} \cdot \frac{1}{f(N)^{j-1} f'(N)} \left[ 1 + O \left( \frac{1}{\ln A} \right) \right] \\ &\quad + O \left( \frac{A^{t-1}}{f(N+1)^j} \right). \end{aligned} \quad (2.28)$$

In view of (2.28) and (2.24), (2.25) yields

$$\begin{aligned} I_N^{12}(\alpha; j) &= \frac{1}{f(N)^{j-1} f'(N) A \ln A} \\ &\quad \times \int_0^\varepsilon \left[ \frac{A^t}{1-t} + \omega(N) \frac{A^t}{(1-t)^2 \ln A} \left[ 1 + O \left( \frac{1}{\ln A} \right) \right] + O(A^{t-1} \ln A) \right] \\ &\quad \times (1-t)^{j-1} \exp \left[ -\frac{A^t}{\ln A} \left( \sum_{n=0}^\infty t^n \right) - \omega(N) \frac{A^t}{\ln^2 A} \left( \sum_{n=1}^\infty n t^{n-1} \right) \right. \\ &\quad \left. \times \left[ 1 + O \left( \frac{1}{\ln A} \right) \right] \right] dt. \end{aligned}$$

Substituting  $u = A^t / \ln A$  in the integral above, we get (in view of (2.24))

$$\begin{aligned}
I_N^{12}(\alpha; j) &= \frac{1}{f(N)^j \ln A} \int_{1/\ln A}^{A^\varepsilon / \ln A} \left\{ \left[ \frac{1}{1 - \frac{\ln u}{\ln A} - \frac{\ln(\ln A)}{\ln A}} \right. \right. \\
&\quad \left. \left. + \frac{\omega(N)}{\ln A} \frac{1}{\left(1 - \frac{\ln u}{\ln A} - \frac{\ln(\ln A)}{\ln A}\right)^2} \left[ 1 + O\left(\frac{1}{\ln A}\right) \right] + O\left(\frac{\ln A}{A}\right) \right] \right. \\
&\quad \left. \times \left( 1 - \frac{\ln u}{\ln A} - \frac{\ln(\ln A)}{\ln A} \right)^{j-1} \right. \\
&\quad \left. \times \exp \left[ -\frac{u}{1 - \frac{\ln u}{\ln A} - \frac{\ln(\ln A)}{\ln A}} - \frac{\omega(N)}{\ln A} \frac{u}{\left(1 - \frac{\ln u}{\ln A} - \frac{\ln(\ln A)}{\ln A}\right)^2} \left[ 1 + O\left(\frac{1}{\ln A}\right) \right] \right] \right\} du.
\end{aligned}$$

If we set

$$\delta := \frac{1}{\ln A} = \frac{1}{\ln \left( \frac{f(N)}{f'(N)} \right)} = \frac{f(N)}{F(N)} \quad (2.29)$$

(hence,  $A \rightarrow \infty$  implies  $\delta \rightarrow 0^+$ ), the above integral becomes

$$\begin{aligned}
I_N^{12}(\alpha; j) &= \frac{\delta}{f(N)^j} \int_{\delta}^{\delta \exp(\varepsilon/\delta)} \left\{ \left[ (1 - \delta \ln u + \delta \ln \delta)^{j-2} \right. \right. \\
&\quad \left. \left. + \omega(N) \delta (1 - \delta \ln u + \delta \ln \delta)^{j-3} (1 + O(\delta)) + O\left(\frac{e^{-1/\delta}}{\delta}\right) \right] \right. \\
&\quad \left. \times \exp \left[ -\frac{u}{1 - \delta \ln u + \delta \ln \delta} - \frac{\omega(N)u\delta}{(1 - \delta \ln u + \delta \ln \delta)^2} (1 + O(\delta)) \right] \right\} du \quad (2.30)
\end{aligned}$$

We split the integral of (2.30) as:

$$\int_{\delta}^{\delta \exp(\varepsilon/\delta)} = \int_{\delta}^{1/\sqrt{\delta}} + \int_{1/\sqrt{\delta}}^{\delta \exp(\varepsilon/\delta)}. \quad (2.31)$$

The second integral of (2.31) can be bounded as follows:

$$\begin{aligned}
&\int_{1/\sqrt{\delta}}^{\delta \exp(\varepsilon/\delta)} \left\{ \left[ (1 - \delta \ln u + \delta \ln \delta)^{j-2} \right. \right. \\
&\quad \left. \left. + \omega(N) \delta (1 - \delta \ln u + \delta \ln \delta)^{j-3} (1 + O(\delta)) + O\left(\frac{e^{-1/\delta}}{\delta}\right) \right] \right. \\
&\quad \left. \times \exp \left[ -\frac{u}{1 - \delta \ln u + \delta \ln \delta} - \frac{\omega(N)u\delta}{(1 - \delta \ln u + \delta \ln \delta)^2} (1 + O(\delta)) \right] \right\} du
\end{aligned}$$

$$\begin{aligned}
&= \int_{1/\sqrt{\delta}}^{\delta \exp(\varepsilon/\delta)} \left\{ (1 - \delta \ln u + \delta \ln \delta)^{j-2} \left[ 1 + \frac{\omega(N)}{1 - \delta \ln u + \delta \ln \delta} \delta (1 + O(\delta)) + O\left(\frac{e^{-1/\delta}}{\delta}\right) \right] \right. \\
&\quad \left. \times \exp \left[ -\frac{u}{1 - \delta \ln u + \delta \ln \delta} \left[ 1 + \frac{\omega(N)}{1 - \delta \ln u + \delta \ln \delta} \delta (1 + O(\delta)) \right] \right] \right\} du \\
&\leq M \int_{1/\sqrt{\delta}}^{\infty} e^{-u} du = O\left(e^{-1/\sqrt{\delta}}\right), \tag{2.32}
\end{aligned}$$

for some positive constant  $M$  (since,  $0 < \varepsilon < 1/2$ ). The first integral of (2.31) is

$$\begin{aligned}
K_1(\delta) &:= \int_{\delta}^{1/\sqrt{\delta}} \left\{ \left[ (1 - \delta \ln u + \delta \ln \delta)^{j-2} \right. \right. \\
&\quad \left. \left. + \omega(N) \delta (1 - \delta \ln u + \delta \ln \delta)^{j-3} (1 + O(\delta)) + O\left(\frac{e^{-1/\delta}}{\delta}\right) \right] \right. \\
&\quad \left. \times \exp \left( -\frac{u}{1 - \delta \ln u + \delta \ln \delta} - \frac{\omega(N)u\delta}{(1 - \delta \ln u + \delta \ln \delta)^2} [1 + O(\delta)] \right) \right\} du
\end{aligned}$$

Since  $(1 - x)^{-2} = \sum_{n=1}^{\infty} nx^{n-1}$  for  $|x| < 1$ , we have

$$\begin{aligned}
K_1(\delta) &= \int_{\delta}^{1/\sqrt{\delta}} \left[ \left( 1 - \delta \ln \frac{u}{\delta} \right)^{j-2} + \omega(N) \delta \left( 1 - \delta \ln \frac{u}{\delta} \right)^{j-3} (1 + O(\delta)) + O\left(\frac{e^{-1/\delta}}{\delta}\right) \right] \\
&\quad \times \exp \left[ -u \sum_{n=0}^{\infty} \left( \delta \ln \frac{u}{\delta} \right)^n - u \omega(N) \delta (1 + O(\delta)) \sum_{n=1}^{\infty} n \left( \delta \ln \frac{u}{\delta} \right)^{n-1} \right] du.
\end{aligned}$$

We use the binomial theorem to expand the quantities  $(1 - \delta \ln \frac{u}{\delta})^{j-2}$  and  $(1 - \delta \ln \frac{u}{\delta})^{j-3}$  and get

$$\begin{aligned}
K_1(\delta) &= \int_{\delta}^{1/\sqrt{\delta}} \left\{ \left[ \left( 1 - (j-2) \delta \ln \frac{u}{\delta} + O\left(\delta^2 \ln^2 \frac{u}{\delta}\right) \right) \right. \right. \\
&\quad \left. \left. + \omega(N) \delta (1 + O(\delta)) \left( 1 + O\left(\delta \ln \frac{u}{\delta}\right) \right) + O\left(\frac{e^{-1/\delta}}{\delta}\right) \right] \right. \\
&\quad \left. \times e^{-u} \exp \left( -u \sum_{n=1}^{\infty} \left( \delta \ln \frac{u}{\delta} \right)^n \right) \exp \left( -u \omega(N) \delta (1 + O(\delta)) \sum_{n=1}^{\infty} n \left( \delta \ln \frac{u}{\delta} \right)^{n-1} \right) \right\} du.
\end{aligned}$$

Next, we expand the exponentials and get (since  $e^x = 1 + x + O(x^2)$  as  $x \rightarrow 0$ )

$$K_1(\delta) = \int_{\delta}^{1/\sqrt{\delta}} \left[ 1 - (j-2) \delta \ln \frac{u}{\delta} + \omega(N) \delta + O\left(\delta^2 \ln^2 \frac{u}{\delta}\right) \right]$$

$$\begin{aligned}
& \times e^{-u} \left\{ 1 - u \sum_{n=1}^{\infty} \left( \delta \ln \frac{u}{\delta} \right)^n + O \left( u \sum_{n=1}^{\infty} \left( \delta \ln \frac{u}{\delta} \right)^n \right)^2 \right\} \\
& \times \left\{ 1 - \omega(N) u \delta (1 + O(\delta)) \sum_{n=1}^{\infty} \left( \delta \ln \frac{u}{\delta} \right)^{n-1} \right. \\
& \left. + O \left( \omega(N) u \delta (1 + O(\delta)) \sum_{n=1}^{\infty} \left( \delta \ln \frac{u}{\delta} \right)^{n-1} \right)^2 \right\} du
\end{aligned}$$

Hence,

$$\begin{aligned}
K_1(\delta) &= \int_{\delta}^{1/\sqrt{\delta}} e^{-u} \left[ 1 - (j-2+u) \delta \ln \frac{u}{\delta} + \omega(N) \delta (1-u) + u^2 O \left( \delta^2 \ln^2 \frac{u}{\delta} \right) \right] du \\
&= \int_{\delta}^{\infty} e^{-u} \left[ 1 - (j-2+u) \delta \ln \frac{u}{\delta} + \omega(N) \delta (1-u) + u^2 O \left( \delta^2 \ln^2 \frac{u}{\delta} \right) \right] du \\
&- \int_{1/\sqrt{\delta}}^{\infty} e^{-u} \left[ 1 - (j-2+u) \delta \ln \frac{u}{\delta} + \omega(N) \delta (1-u) + u^2 O \left( \delta^2 \ln^2 \frac{u}{\delta} \right) \right] du.
\end{aligned} \tag{2.33}$$

However,

$$\begin{aligned}
& \int_{1/\sqrt{\delta}}^{\infty} e^{-u} \left[ 1 - (j-2+u) \delta \ln \frac{u}{\delta} + \omega(N) \delta (1-u) + u^2 O \left( \delta^2 \ln^2 \frac{u}{\delta} \right) \right] du \\
&= \int_{1/\sqrt{\delta}}^{\infty} e^{-u} \left[ 1 - u \left( \delta \ln \frac{u}{\delta} + \omega(N) \delta \right) + \omega(N) \delta - (j-2) \delta \ln \frac{u}{\delta} \right] du \\
&\leq \int_{1/\sqrt{\delta}}^{\infty} e^{-u} \left[ 1 - \left( 1/\sqrt{\delta} \right) \left( \delta \ln \left( \frac{1/\sqrt{\delta}}{\delta} \right) + \omega(N) \delta \right) + \omega(N) \delta - (j-2) \delta \ln \left( \frac{1/\sqrt{\delta}}{\delta} \right) \right] du \\
&\quad + \int_{1/\sqrt{\delta}}^{\infty} u^2 e^{-u} O \left( \delta^2 \ln^2 \frac{u}{\delta} \right) du = O \left( e^{-1/\sqrt{\delta}} \right) \quad \text{as } \delta \rightarrow 0^+.
\end{aligned} \tag{2.34}$$

It follows that in the expression for  $K_1(\delta)$  we can replace the upper limit of the integral by  $\infty$ . Therefore, (2.30) becomes (as  $\delta \rightarrow 0^+$ )

$$\begin{aligned}
I_N^{12}(\alpha; j) &= \frac{\delta}{f(N)^j} \int_{\delta}^{\infty} e^{-u} \left[ 1 - (j-2+u) \delta \ln \frac{u}{\delta} + \omega(N) \delta (1-u) + u^2 O \left( \delta^2 \ln^2 \frac{u}{\delta} \right) \right] du \\
&= \frac{\delta}{f(N)^j} \left\{ [1 + (j-2) \delta \ln \delta + \omega(N) \delta] e^{-\delta} + \delta [\ln \delta - \omega(N)] (1+\delta) e^{-\delta} \right\} \\
&- \frac{\delta}{f(N)^j} \left\{ \delta \int_{\delta}^{\infty} e^{-u} u \ln u du + (j-2) \delta \int_{\delta}^{\infty} e^{-u} \ln u du \right\} \\
&+ \frac{\delta}{f(N)^j} O \left( \int_{\delta}^{\infty} e^{-u} u^2 \delta^2 \ln^2 \frac{u}{\delta} du \right).
\end{aligned} \tag{2.35}$$



To continue we need some lemmas.

**Lemma 2.** For the integral,

$$M(x) := \int_x^\infty e^{-t} t \ln t \, dt,$$

we have the asymptotic expansion, as  $x \rightarrow 0^+$ ,

$$M(x) \sim 1 - \gamma + \frac{x^2}{2} \ln x - \frac{x^2}{4} + \dots \quad (2.36)$$

*Proof.* Since

$$\frac{dM(x)}{dx} = -x \ln x e^{-x} = -x \ln x \left( 1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3 + \dots \right),$$

we have

$$M(x) \sim C_1 - \frac{1}{2}x^2 \ln x + \frac{1}{4}x^2 - \frac{1}{3}x^3 \ln x - \frac{1}{9}x^3 - \frac{1}{8}x^4 \ln x + \frac{1}{32}x^4 + \dots, \quad (2.37)$$

where  $C_1$  is a constant. In fact,

$$C_1 = \int_0^\infty t \ln t e^{-t} dt.$$

Integration by parts yields

$$C_1 = \int_0^\infty e^{-t} \ln t \, dt + \int_0^\infty e^{-t} dt = \Gamma'(1) + 1 = 1 - \gamma$$

(see [6]) and the proof is completed. ■

**Lemma 3.** For the integral,

$$G(x) := \int_x^\infty \ln t e^{-t} dt,$$

we have the asymptotic expansion, as  $x \rightarrow 0^+$ ,

$$G(x) \sim -\gamma - x \ln x + x + \frac{1}{2}x^2 \ln x - \frac{1}{2}x^2 - \frac{1}{6}x^3 \ln x + \frac{1}{6}x^3 + \frac{1}{24}x^4 \ln x - \frac{1}{24}x^4 + \dots \quad (2.38)$$

The proof is similar to Lemma 2; it has been given in [9].

We also observe that as  $x \rightarrow 0^+$ ,

$$\int_x^\infty t^2 e^{-t} \ln^2 t \, dt \sim C_0 := \int_0^\infty t^2 e^{-t} \ln^2 t \, dt = \frac{\pi^2}{3} + 2\gamma^2 - 6\gamma + 2,$$

and

$$\int_x^\infty t^2 e^{-t} \ln t \, dt \sim \tilde{C}_0 := \int_0^\infty t^2 e^{-t} \ln t \, dt = 3 - 2\gamma.$$

In particular,

$$O\left(\int_\delta^\infty e^{-u} u^2 \delta^2 \ln^2 \frac{u}{\delta} \, du\right) = O(\delta^2 \ln^2 \delta) \quad \text{as } \delta \rightarrow 0^+. \quad (2.39)$$

Applying Lemmas 2, 3, and (2.39), in (2.35) we get (since  $e^{-\delta} = 1 - \delta + O(\delta^2)$  as  $\delta \rightarrow 0^+$ ),

$$I_N^{12}(\alpha; j) = \frac{\delta}{f(N)^j} [1 + (j-1)\delta \ln \delta + [(j-1)\gamma - 2]\delta + O(\delta^2 \ln^2 \delta)]. \quad (2.40)$$

Notice that the error term in (2.40) dominates the terms of (2.32) and (2.34).

**Remark 5.** In view of (2.29), (2.20) yields

$$I_N^{11}(\alpha; j) << \frac{\delta^3 \ln^2 \delta}{f(N)^j}, \quad j \geq 2, \quad (2.41)$$

as  $N \rightarrow \infty$ . Using (2.41) and (2.40) and invoking (2.12), one has

$$I_N^1(\alpha; j) = \frac{\delta}{f(N)^j} [1 + (j-1)\delta \ln \delta + [(j-1)\gamma - 2]\delta + O(\delta^2 \ln^2 \delta)]. \quad (2.42)$$

## 2.2 The integral $I_N^2(\alpha; j)$

Our next task is to compute the asymptotic behavior of the quantity  $I_N^2(\alpha; j)$  defined in (2.8). It has been established in [5] that,

$$\lim_N \sum_{i=1}^N \ln \left(1 - e^{-\frac{F(N)}{f(i)}s}\right) = 0,$$

uniformly in  $s \in [1, \infty)$ . From (2.8) we have

$$I_N^2(\alpha; j) = \sum_{k=1}^N f(k)^{-j} \int_1^\infty s^{j-1} \frac{e^{-\frac{F(N)}{f(k)}s}}{1 - e^{-\frac{F(N)}{f(k)}s}} \exp \left[ \sum_{i=1}^N \ln \left(1 - e^{-\frac{F(N)}{f(i)}s}\right) \right] ds. \quad (2.43)$$

Using the Taylor expansion of the logarithm, the comparison of sums and integrals (i.e. (2.16)), and Lemma 1, for  $m = 0$ , we get

$$I_N^2(\alpha; j) = \sum_{k=1}^N f(k)^{-j} \int_1^\infty s^{j-1} \frac{e^{-\frac{F(N)}{f(k)}s}}{1 - e^{-\frac{F(N)}{f(k)}s}} \left[ 1 + O\left(\frac{f(N)^2}{F(N)f'(N)} e^{-\frac{F(N)}{f(N)}s}\right) \right] ds.$$

The scaling  $u = F(N)f(k)^{-1}s$ , via the definition of  $F$  yields

$$\begin{aligned} \int_1^\infty s^{j-1} \frac{e^{-\frac{F(N)}{f(k)}s}}{1 - e^{-\frac{F(N)}{f(k)}s}} ds &= \left( \frac{f(k)}{F(N)} \right)^j \int_{\frac{F(N)}{f(k)}}^\infty \frac{e^{-u}}{1 - e^{-u}} u^{j-1} du \\ &= \frac{f(k)}{F(N)} e^{-\frac{F(N)}{f(k)}} + O \left[ \left( \frac{f(k)}{F(N)} \right)^2 e^{-\frac{F(N)}{f(k)}} \right]. \end{aligned} \quad (2.44)$$

Thus,  $I_N^2(\alpha; j)$  (see (2.43)), becomes

$$\begin{aligned} I_N^2(\alpha; j) &= \sum_{k=1}^N f(k)^{-j} \left( \frac{f(k)}{F(N)} e^{-\frac{F(N)}{f(k)}} + O \left[ \left( \frac{f(k)}{F(N)} \right)^2 e^{-\frac{F(N)}{f(k)}} \right] \right) \\ &= \frac{1}{F(N)} \sum_{k=1}^N \frac{1}{f(k)^{j-1}} e^{-\frac{F(N)}{f(k)}} + O \left[ \frac{1}{F(N)^2} \sum_{k=1}^N \frac{1}{f(k)^{j-2}} e^{-\frac{F(N)}{f(k)}} \right]. \end{aligned}$$

Under conditions (2.3), for sufficiently large  $N$  we have

$$\sum_{k=1}^N \frac{1}{f(k)^{j-1}} e^{-\frac{F(N)}{f(k)}} = \int_1^N \frac{1}{f(x)^{j-1}} e^{-\frac{F(N)}{f(x)}} dx + O \left( \frac{1}{f(N+1)^{j-1}} e^{-\frac{F(N)}{f(N+1)}} \right).$$

Applying Lemma 1, for  $m = (1-j)$ ,  $s = 1$ , yields

$$\begin{aligned} \sum_{k=1}^N \frac{1}{f(k)^{j-1}} e^{-\frac{F(N)}{f(k)}} &= \frac{f(N)^{3-j}}{F(N)f'(N)} e^{-\frac{F(N)}{f(N)}} + \omega(N) \frac{f(N)^{4-j}}{F(N)^2 f'(N)} e^{-\frac{F(N)}{f(N)}} \left[ 1 + O \left( \frac{f(N)}{F(N)} \right) \right] \\ &\quad + O \left( \frac{1}{f(N+1)^{j-1}} e^{-\frac{F(N)}{f(N+1)}} \right). \end{aligned}$$

By (2.4) the above quantity becomes

$$\sum_{k=1}^N \frac{1}{f(k)^{j-1}} e^{-\frac{F(N)}{f(k)}} = \frac{f(N)^{3-j}}{F(N)f'(N)} e^{-\frac{F(N)}{f(N)}s} + \omega(N) \frac{f(N)^{4-j}}{F(N)^2 f'(N)} e^{-\frac{F(N)}{f(N)}s} \left[ 1 + O \left( \frac{f(N)}{F(N)} \right) \right].$$

Hence,

$$I_N^2(\alpha; j) = \frac{f(N)^{3-j}}{F(N)^2 f'(N)} e^{-\frac{F(N)}{f(N)}} \left[ 1 + O \left( \frac{f(N)}{F(N)} \right) \right].$$

Using the definition of  $F$ , i.e. (2.5), and (2.29) one has

$$I_N^2(\alpha; j) = \frac{1}{f(N)^j \ln^j \left( \frac{f(N)}{f'(N)} \right)} \left[ 1 + O \left( \frac{1}{\ln \left( \frac{f(N)}{f'(N)} \right)} \right) \right] = \frac{\delta}{f(N)^j} (\delta^{j-1} + O(\delta^j)). \quad (2.45)$$

We are now ready for our main result.

### 2.3 Conclusion. Asymptotics of $E[U_j^N]$

Recall that when the main collector has completed her album, they remain  $U_j^N$  unfilled places in the album of the  $j$ -th collector,  $j = 2, 3, \dots$ . The asymptotics of  $E[U_j^N]$  is given by the following theorem.

**Theorem 1.** Let  $\alpha = \{a_k\}_{k=1}^\infty = \{1/f(k)\}_{k=1}^\infty$ , where  $f$  satisfies (2.2) and (2.3). If the coupon probabilities (i.e. the  $p_k$ 's) are as in (1.10), then, as  $N \rightarrow \infty$ , we have

$$E[U_2^N] = \frac{1}{\delta} [1 + \delta \ln \delta + (\gamma - 1) \delta + O(\delta^2 \ln^2 \delta)], \quad (2.46)$$

$$E[U_j^N] = \frac{1}{(j-1)! \delta^{j-1}} [1 + (j-1) \delta \ln \delta + [(j-1) \gamma - 2] \delta + O(\delta^2 \ln^2 \delta)] \quad (2.47)$$

for  $j \geq 3$ . Recall that

$$\delta = \frac{1}{\ln(f(N)/f'(N))}.$$

*Proof.* The desired result follows by using formulas (2.42) and (2.45) in (2.6) and (1.19). Notice that, for  $j \geq 3$  all three terms of the asymptotics in (2.47) come solely from  $I_N^{12}(\alpha; j)$  (see (2.40)). For  $j = 2$  part of the third term of the asymptotics in (2.46) is due to  $I_N^2(\alpha; j)$  (see (2.45)), while the rest of the third term, as well as the first two terms are, again, due to  $I_N^{12}(\alpha; j)$  (see (2.40)). The integral  $I_N^{12}(\alpha; j)$  does not contribute at all in the first three terms of the asymptotics of  $E[U_j^N]$ . ■

**Remark 6.** From Theorem 1 and for all  $j \geq 2$  we have

$$E[U_j^N] \sim \frac{1}{(j-1)!} \ln \left( \frac{f(N)}{f'(N)} \right)^{j-1}, \quad N \rightarrow \infty. \quad (2.48)$$

**Example 1.**  $a_k = 1/k^p$ , where  $p > 0$ . This is the so-called *generalized Zipf* law (for detailed asymptotic results regarding the first collector, i.e. the random variable  $T_N$ , see [9]). These decaying sequences fall, clearly, into the previous discussion, since  $f(x) = x^p$ , satisfies (i)–(iv) of (2.3). Here,

$$\delta = \frac{1}{\ln N - \ln p} \quad (2.49)$$

and, hence, Theorem 1 gives

$$E[U_2^N] = \ln N - \ln(\ln N) + (\gamma - 1 - \ln p) + O\left(\frac{\ln(\ln N)^2}{\ln N}\right) \quad (2.50)$$

and

$$E[U_j^N] = \frac{(\ln N)^{j-1}}{(j-1)!} - \frac{(\ln N)^{j-2} \ln(\ln N)}{(j-2)!} + \frac{(j-1)(\gamma - \ln p) - 2}{(j-1)!} (\ln N)^{j-2} + O\left((\ln N)^{j-3} \ln(\ln N)^2\right) \quad (2.51)$$

for  $j \geq 3$ . Notice that  $p$  does not appear in the first two terms of the asymptotics. Also, the leading term is the same as in the case where all the  $p_k$ 's are equal (see (1.5)), however the second terms of the asymptotics differ.

**Example 2.**  $a_k = \exp(-pk^q)$ , where  $p > 0$  and  $0 < q < 1$ . Again, these decaying sequences fall into the previous discussion, since  $f(x) = \exp(px^q)$ , satisfies (i)–(iv) of (2.3). Here,

$$\delta = \frac{1}{\ln(N^{1-q}) - \ln(pq)}.$$

If we compare the above formula with (2.49) we can see that the asymptotics of  $E[U_j^N]$  ( $j \geq 2$ ) for Example 2 can be obtained from the formulas (2.50) and (2.51) of Example 1 after replacing  $N$  by  $N^{1-q}$  and  $p$  by  $pq$ . For example (this, also follows from (2.48)),

$$E[U_j^N] \sim \frac{(\ln(N^{1-q}))^{j-1}}{(j-1)!} = (1-q)^{j-1} \frac{(\ln N)^{j-1}}{(j-1)!}, \quad N \rightarrow \infty,$$

for all  $j \geq 2$ . Notice that  $p$  does not appear in the leading asymptotics of  $E[U_j^N]$ ; it first appears in the third term.

### 3 Growing sequences

In this section we will examine sequences  $\alpha = \{a_k\}_{k=1}^\infty$ , such that

$$a_k \rightarrow \infty.$$

We will exhibit several cases where (recall (1.19)),  $E[U_j^N] = I_N(\alpha; j)$  approaches a finite limit as  $N \rightarrow \infty$ . However, we will also see that there are cases for which  $\lim_N E[U_j^N] = \infty$ .

By substituting  $x = e^{-t}$ , (1.17) becomes

$$I_N(\alpha; j) = \frac{1}{(j-1)!} \int_0^1 \left( \sum_{k=1}^N a_k^j \frac{x^{a_k}}{1-x^{a_k}} \right) \left[ \prod_{k=1}^N (1-x^{a_k}) \right] |\ln x|^{j-1} \frac{dx}{x}. \quad (3.1)$$

For typographical convenience let us set

$$S_N(x) = S_N(x; \alpha) := \sum_{k=1}^N x^{a_k}, \quad (3.2)$$

$$F_N(x) = F_N(x; \alpha) := \prod_{k=1}^N (1-x^{a_k}), \quad (3.3)$$

and

$$L_N(x; j) = L_N(x; \alpha; j) := \sum_{k=1}^N a_k^j \frac{x^{a_k}}{1 - x^{a_k}}, \quad j = 2, \dots, r \quad (3.4)$$

(in fact,  $L_N(x; \alpha; j)$  makes sense for any real number  $j$ ). Obviously, for a fixed  $x \in (0, 1)$  and a fixed  $j$  we have that  $S_N(x)$  and  $L_N(x; j)$  increase with  $N$ , while  $F_N(x)$  decreases.

Given a sequence  $\alpha = \{a_k\}_{k=1}^\infty$  of positive terms let

$$x_\alpha := \inf \left\{ x \in [0, 1] : \sum_{k=1}^\infty x^{a_k} = \infty \right\}. \quad (3.5)$$

From now on, we will consider only sequences  $\alpha$  such that

$$0 < x_\alpha \leq 1. \quad (3.6)$$

Roughly speaking, condition (3.5)–(3.6) says that  $a_k$  grows at least logarithmically. For example, if  $a_k = \ln k$  ( $k \geq 2$ ), then  $x_\alpha = 1/e$  and (3.6) is satisfied. However, if  $a_k = \epsilon_k \ln k$ , where  $\epsilon_k \rightarrow 0$ , then  $x_\alpha = 0$ , i.e. (3.6) is not satisfied. To have  $x_\alpha = 1$ ,  $a_k$  must grow faster than  $\ln k$  (roughly speaking). For instance, if  $a_k = \lambda_k \ln k$ , where  $\lambda_k \rightarrow \infty$ , then  $x_\alpha = 1$ .

Let us set

$$S(x) = S(x; \alpha) := \sum_{k=1}^\infty x^{a_k}. \quad (3.7)$$

Then, from condition (3.5)–(3.6) it follows that  $S(x) < \infty$  for  $x \in [0, x_\alpha)$  and  $S(x) = \infty$  for  $x \in (x_\alpha, 1]$ . Furthermore, if  $x_\alpha < 1$ , then, depending on  $\alpha$ ,  $S(x_\alpha)$  can be finite or infinite. For instance, if  $a_k = \ln k$  ( $k \geq 2$ ), then  $x_\alpha = 1/e$  and  $S(1/e) = \infty$ , while if  $a_k = \ln(k \ln^2 k)$  ( $k \geq 3$ ), then, again  $x_\alpha = 1/e$ , but now  $S(1/e)$  is finite. Of course, if  $x_\alpha = 1$ , then  $S(x_\alpha) = S(1) = \infty$ .

If we set

$$F(x) = F(x; \alpha) := \lim_N F_N(x; \alpha) = \prod_{k=1}^\infty (1 - x^{a_k}), \quad (3.8)$$

then, in view of (3.5)–(3.6), by standard properties of infinite products (see, e.g., [27]) we have that

$$0 < F(x) < 1 \quad \text{for all } x \in (0, x_\alpha) \quad (3.9)$$

and, also,  $F(x) = 0$  for  $x \in (x_\alpha, 1]$ . Of course,  $F(0) = 1$  and  $F(1) = 0$ . Furthermore,  $F(x)$  is (decreasing on  $[0, 1]$  and) continuous for all  $x \in [0, 1]$  with only one possible exception (at  $x = x_\alpha$ ):  $F(x)$  is not continuous at  $x_\alpha$  if and only if  $S(x_\alpha) < \infty$ , since in this case  $F(x_\alpha-) = F(x_\alpha) > 0$ , while

$F(x_\alpha+) = 0$ . Of course, if  $x_\alpha = 1$ , then  $F(x)$  is continuous on  $[0, 1]$ .  
Next, we notice that it is easy to show that, under (3.5)–(3.6) we have

$$L(x; j) = L(x; \alpha; j) := \lim_N L_N(x; \alpha; j) = \sum_{k=1}^{\infty} a_k^j \frac{x^{a_k}}{1 - x^{a_k}} < \infty, \quad x \in (0, x_\alpha). \quad (3.10)$$

Furthermore,

$$L(x; j) = O(x^{a_{\min}}) \quad \text{as } x \rightarrow 0^+, \quad (3.11)$$

where

$$a_{\min} := \min_{k \in \mathbb{N}} (a_k) > 0 \quad (3.12)$$

(the strict positivity follows from the fact that  $a_k > 0$  for all  $k \in \mathbb{N}$  and  $a_k \rightarrow \infty$ ). As for the value of  $L(x_\alpha; j)$ , depending on the sequence  $\alpha$  there might be a  $j_0$  such that  $L(x_\alpha; j) < \infty$ , for  $j < j_0$ , while  $L(x_\alpha; j) = \infty$ , for  $j \geq j_0$ . For example, if  $a_k = \ln(k \ln^p k)$  ( $k \geq 3$ ), then  $x_\alpha = 1/e$  and  $L(1/e; j) < \infty$  for  $j < p - 1$ , while  $L(1/e; j) = \infty$  for  $j \geq p - 1$ .

Inspired by (3.1) we introduce the quantity

$$I(\alpha; j) := \frac{1}{(j-1)!} \int_0^{x_\alpha} L(x; \alpha; j) F(x; \alpha) |\ln x|^{j-1} \frac{dx}{x} \quad (3.13)$$

$$= \frac{1}{(j-1)!} \int_{-\ln(x_\alpha)}^{\infty} L(e^{-t}; \alpha; j) F(e^{-t}; \alpha) t^{j-1} dt. \quad (3.14)$$

**Proposition 2.** If  $I(\alpha; j) = \infty$  (see (3.13)), then

$$\lim_N E[U_j^N] = \infty.$$

*Proof.* First notice that by (1.19), (3.1), (3.3), and (3.4) we have

$$\begin{aligned} E[U_j^N] &= I_N(\alpha; j) = \frac{1}{(j-1)!} \int_0^1 L_N(x; \alpha; j) F_N(x; \alpha) |\ln x|^{j-1} \frac{dx}{x} \\ &\geq \frac{1}{(j-1)!} \int_0^{x_\alpha} L_N(x; \alpha; j) F_N(x; \alpha) |\ln x|^{j-1} \frac{dx}{x}. \end{aligned}$$

Also, by (3.8) and (3.10)

$$\lim_N [L_N(x; \alpha; j) F_N(x; \alpha)] = L(x; \alpha; j) F(x; \alpha), \quad 0 \leq x < x_\alpha.$$

Thus, the proposition follows from (3.13) and the Fatou Lemma. ■

### 3.1 Some results for the case $x_\alpha = 1$

**Proposition 3.** Assume that for the sequence  $\alpha = \{a_k\}_{k=1}^\infty$  we have  $x_\alpha = 1$  (recall (3.5)). If there is an  $N_0$  such that

$$\int_0^1 L(x; \alpha; j) F_{N_0}(x; \alpha) |\ln x|^{j-1} \frac{dx}{x} < \infty, \quad (3.15)$$

then (recalling (3.13))

$$\lim_N E[U_j^N] = I(\alpha; j) < \infty. \quad (3.16)$$

*Proof.* As in the proof of Proposition 2

$$E[U_j^N] = I_N(\alpha; j) = \frac{1}{(j-1)!} \int_0^1 L_N(x; \alpha; j) F_N(x; \alpha) |\ln x|^{j-1} \frac{dx}{x}.$$

and

$$\lim_N [L_N(x; \alpha; j) F_N(x; \alpha)] = L(x; \alpha; j) F(x; \alpha), \quad 0 \leq x < 1.$$

Since

$$0 < L_N(x; \alpha; j) F_N(x; \alpha) < L(x; \alpha; j) F_{N_0}(x; \alpha) \quad \text{for } N > N_0.$$

(3.16) follows by dominated convergence. ■

If  $\alpha = \{a_k\}_{k=1}^\infty$  (with  $x_\alpha = 1$ ) satisfies  $a_k = o(k^p)$  for all  $p > 0$  then condition (3.15) cannot be satisfied. As an example of such a sequence one can take  $a_k = (\ln k)^q$ ,  $q > 1$ . In this case Proposition 3 is inconclusive.

**Remark 7.** Suppose that for the sequence  $\alpha = \{a_k\}_{k=1}^\infty$  we have  $x_\alpha = 1$ . Let  $\beta = \{b_k\}_{k=1}^\infty$  be a sequence such that there is an integer  $k_0 \geq 0$  for which

$$b_k = a_{k+k_0}, \quad k \in \mathbb{N} \quad (3.17)$$

(i.e.  $\beta$  is the  $k_0$ -left shift of  $\alpha$ ). Then  $x_\beta = 1$  and it is easy to see that  $\alpha$  satisfies condition (3.15) for some  $N_0 = N_0(\alpha)$  if and only if  $\beta$  satisfies condition (3.15) for some  $N_0 = N_0(\beta)$ .

The same equivalence is true if given  $\alpha$  the sequence  $\tilde{\beta} = \{\tilde{b}_k\}_{k=1}^\infty$  is such that

$$|a_k - \tilde{b}_k| \leq M, \quad k \in \mathbb{N}, \quad (3.18)$$

for some  $M > 0$ . In particular, for a given sequence  $\alpha = \{a_k\}_{k=1}^\infty$ , if we set  $\underline{\alpha} := \{\lfloor a_k \rfloor\}_{k=1}^\infty$  and  $\overline{\alpha} := \{\lceil a_k \rceil\}_{k=1}^\infty$  (where  $\lfloor x \rfloor$  and  $\lceil x \rceil$  denote the greatest integer  $\leq x$  and the smallest integer  $\geq x$  respectively), then  $\alpha$  satisfies condition (3.15) if and only if  $\underline{\alpha}$  satisfies condition (3.15) (and this in turn



holds if and only if  $\bar{\alpha}$  satisfies condition (3.15)).

Let us now assume that all the terms of the sequence  $\alpha = \{a_k\}_{k=1}^{\infty}$  are positive *integers*. Then  $L(x; \alpha; j)$  of (3.10) can be expressed as

$$L(x; \alpha; j) = \sum_{m=1}^{\infty} m^j A(m) \frac{x^m}{1 - x^m} \quad (3.19)$$

where

$$A(m) := \# \{a_k : a_k = m\} \quad (3.20)$$

(the symbol  $\#$  indicates cardinality). The series in the right-hand side of (3.19) is a so-called Lambert series (see, e.g., [17]) and can be easily transformed to a power series. Indeed,

$$L(x; \alpha; j) = \sum_{n=1}^{\infty} A_L(n) x^n, \quad (3.21)$$

where

$$A_L(n) := \sum_{d|n} d^j A(d), \quad (3.22)$$

i.e. the sum is taken over all divisors  $d$  of  $n$ .

**Corollary 1.** Suppose that the sequence  $\alpha = \{a_k\}_{k=1}^{\infty}$  has integer terms and satisfies (recall (3.10), (3.20), (3.21), and (3.22))

$$L(x; \alpha; j) = \sum_{n=1}^{\infty} A_L(n) x^n = O\left(\frac{1}{(1-x)^{\rho}}\right), \quad x \rightarrow 1^-, \quad (3.23)$$

for some  $\rho > 0$ . Then  $\alpha$  satisfies (3.15) and, consequently,

$$\lim_N E[U_j^N] = I(\alpha; j) < \infty.$$

*Proof.* If (3.23) is true, then by choosing  $N_0 = \lceil \rho \rceil$  we get that (see (3.3))

$$L(x; \alpha; j) F_{N_0}(x; \alpha) \quad \text{is bounded for } x \in [0, 1),$$

hence (3.15) is satisfied (with the help of (3.11), which takes care of the lower limit, i.e.  $x = 0$ , of the integral in (3.15)).  $\blacksquare$

**Corollary 2.** Let  $\alpha = \{a_k\}_{k=1}^{\infty}$  be a sequence of positive integers for which  $A(m)$  of (3.20) satisfies

$$A(m) = O(m^{\nu}) \quad (3.24)$$

for some  $\nu > 0$ . Then

$$\lim_N E[U_j^N] = I(\alpha; j) < \infty \quad \text{for all } j \geq 2.$$

*Proof.* From (3.22) and (3.24) we get that

$$A_L(n) = \sum_{d|n} d^j A(d) \leq \sum_{d=1}^n d^j A(d) = O(n^{\nu+j+1}), \quad (3.25)$$

Set  $\mu = \lceil \nu + j + 1 \rceil$ . Then, for  $x \in [0, 1)$  condition (3.25) implies

$$L(x; \alpha; j) \leq M \sum_{n=0}^{\infty} (n+1)(n+2) \cdots (n+\mu) x^n = M \frac{d^\mu}{dx^\mu} \left[ \frac{1}{1-x} \right]$$

for some constant  $M > 0$ . In other words

$$L(x; \alpha; j) \leq \frac{M\mu!}{(1-x)^{\mu+1}}$$

and hence (3.23) of Corollary 1 is satisfied by choosing  $\rho = \mu + 1$ . ■

**Remark 8.** Suppose that  $\alpha = \{a_k\}_{k=1}^{\infty}$  is a sequence whose terms are positive reals, not necessarily integers. Let

$$A^*(m) := \# \{a_k : a_k \leq m\}. \quad (3.26)$$

Notice that

$$A^*(m) \geq \# \{a_k : \lceil a_k \rceil \leq m\}. \quad (3.27)$$

If there is a  $\nu > 0$  such that

$$A^*(m) = O(m^\nu) \quad (3.28)$$

then (3.27) implies that

$$\# \{a_k : \lceil a_k \rceil \leq m\} = O(m^\nu),$$

i.e.  $\bar{\alpha} = \{\lceil a_k \rceil\}_{k=1}^{\infty}$  satisfies (3.24). Therefore, with the help of the last part of Remark 7 we deduce that  $\lim_N E[U_j^N] = I(\alpha; j) < \infty$  for all  $j \geq 2$ .

**Example 3.** Let  $a_k = k^p$ , where  $p > 0$  (the case  $p = 1$  is known as the *linear* case). Then (recall (3.26))

$$A^*(m) = \# \{k \in \mathbb{N} : k^p \leq m\} = O(m^{1/p}).$$

Thus, by Remark 8 we obtain that

$$\lim_N E[U_j^N] = I(\alpha; j) < \infty, \quad \text{for all } j \geq 2.$$

**Example 4.** Let  $a_k = e^{pk^q}$ , where  $p, q > 0$ . Then, we can, again, use Remark 8 as in Example 3 to conclude that

$$\lim_N E[U_j^N] = I(\alpha; j) < \infty, \quad \text{for all } j \geq 2.$$

In the same way we can see that for the sequence  $a_k = k!$ ,  $k = 1, 2, \dots$ , we also have  $\lim_N E[U_j^N] = I(\alpha; j) < \infty$  for all  $j \geq 2$ .

Let us, also, discuss the sequence  $\beta = \{b_k\}_{k=1}^\infty$ , with  $b_k = e^{-pk}$ ,  $p > 0$ . In this case, the function  $f(x) = e^{px}$  does not satisfy condition (ii) of (2.3), thus Theorem 1 cannot be applied. However, the sequences  $\alpha = \{e^{pk}\}_{k=1}^\infty$  and  $\beta$  produce the same coupon probabilities! This follows from the fact that for each  $N$ , if we let  $c_N = e^{pN}$ , then  $\{a_j : 0 \leq j \leq N\} = \{c_N b_j : 0 \leq j \leq N\}$ , i.e. the elements of the two truncated sequences are proportional to each others. It follows that

$$\lim_N E[U_j^N] = I(\alpha; j) < \infty, \quad \text{for all } j \geq 2.$$

Here,  $I(\alpha; j)$  depends on  $p$  (compare with Example 2).

### 3.2 Two examples for the case $0 < x_\alpha < 1$

**Example 5.** If  $a_k = \ln(k \ln^2 k) = \ln k + 2 \ln(\ln k)$  ( $k \geq 3$ ), then

$$\lim_N E[U_j^N] = \infty \quad \text{for all } j \geq 2.$$

In order to justify this equation let us observe that, by Proposition 2 it suffices to show that  $I(\alpha; j) = \infty$ , where  $I(\alpha; j)$  is defined in (3.13)–(3.14). First we notice that  $x_\alpha = 1/e$  and  $S(1/e) < \infty$ .

Clearly,  $1 > F(e^{-t}) \geq F(1/e) > 0$ . Thus, from (3.14) we have

$$I(\alpha; j) = \frac{1}{(j-1)!} \int_1^\infty t^{j-1} \left( \sum_{k=3}^\infty \frac{a_k^j e^{-a_k t}}{1 - e^{-a_k t}} \right) F(e^{-t}) dt.$$

Since  $1/(1 - e^{-a_k t}) > 1$  for all  $t > 0$ , we have (with the help of Tonelli's theorem)

$$\begin{aligned} I(\alpha; j) &\geq \frac{F(1/e)}{(j-1)!} \int_1^\infty t^{j-1} \left( \sum_{k=3}^\infty a_k^j e^{-a_k t} \right) dt \\ &\geq \frac{F(1/e)}{(j-1)!} \int_1^\infty \left( \sum_{k=3}^\infty a_k^j e^{-a_k t} \right) dt \\ &= \frac{F(1/e)}{(j-1)!} \sum_{k=3}^\infty a_k^j \left( \int_1^\infty e^{-a_k t} \right) dt \\ &= \frac{F(1/e)}{(j-1)!} \sum_{k=3}^\infty a_k^{j-1} e^{-a_k}, \end{aligned}$$

and substituting  $a_k = \ln(k \ln^2 k)$  we get

$$I(\alpha; j) \geq \frac{F(1/e)}{(j-1)!} \sum_{k=3}^{\infty} \frac{[\ln(k \ln^2 k)]^{j-1}}{k \ln^2 k} = \infty, \quad j \geq 2.$$

Finally, it is instructive to compare Example 5 with the following example.

**Example 6.** Let  $a_k = \ln k$  ( $k \geq 3$ ). Then,  $x_\alpha = 1/e$  and  $S(1/e) = \infty$  (equivalently,  $S(e^{-t}) = \sum_{k=2}^{\infty} e^{-ta_k} < \infty$  if and only if  $t > 1$ ). Here, in contrast with Example 5, we will show that, for all  $j \geq 2$ ,  $E[U_j^N]$  approaches a finite limit as  $N \rightarrow \infty$ .

From (1.19) and (1.17) we have

$$E[U_j^N] = \frac{1}{(j-1)!} [Q_N^1(\alpha; j) + Q_N^\infty(\alpha; j)], \quad (3.29)$$

where

$$Q_N^1(\alpha; j) := \int_0^1 \left[ \sum_{k=3}^N a_k^j \frac{e^{-a_k t}}{1 - e^{-a_k t}} t^{j-1} \right] \left[ \prod_{k=3}^N (1 - e^{-a_k t}) \right] dt \quad (3.30)$$

and

$$Q_N^\infty(\alpha; j) := \int_1^\infty \left[ \sum_{k=3}^N a_k^j \frac{e^{-a_k t}}{1 - e^{-a_k t}} t^{j-1} \right] \left[ \prod_{k=3}^N (1 - e^{-a_k t}) \right] dt. \quad (3.31)$$

Notice that  $a_k = \ln k$  yields  $e^{-a_k t} = k^{-t}$ . Let us analyze  $Q_N^1(\alpha; j)$  first. Since

$$\frac{x}{1 - e^{-x}} < 1 + x, \quad x \in (0, \infty),$$

formula (3.30) implies

$$Q_N^1(\alpha; j) < \int_0^1 \left[ \sum_{k=3}^N a_k^{j-1} e^{-a_k t} (1 + a_k t) t^{j-2} \right] \left[ \prod_{k=3}^N (1 - e^{-a_k t}) \right] dt$$

or (since  $a_k = \ln k$ )

$$\begin{aligned} Q_N^1(\alpha; j) &< \int_0^1 \left[ \sum_{k=3}^N \left( \frac{(\ln k)^{j-1}}{k^t} t^{j-2} + \frac{(\ln k)^j}{k^t} t^{j-1} \right) \right] \left[ \prod_{k=3}^N \left( 1 - \frac{1}{k^t} \right) \right] dt \\ &< 2 \int_0^1 \left[ \sum_{k=3}^N \frac{(\ln k)^j}{k^t} \right] t^{j-2} \left[ \prod_{k=3}^N \left( 1 - \frac{1}{k^t} \right) \right] dt. \end{aligned} \quad (3.32)$$

The expression  $(\ln k)^j / k^t$  (viewed as a function of  $k$ ) has a unique maximum. It is attained when  $\ln k = j/t$  and the maximum value is  $j^j e^{-j} / t^j$ . It follows that

$$\sum_{k=3}^N \frac{(\ln k)^j}{k^t} < \frac{j^j e^{-j}}{t^j} + \int_3^N \frac{(\ln \kappa)^j}{\kappa^t} d\kappa < \frac{j^j e^{-j}}{t^j} + \int_0^{\ln N} \xi^j e^{(1-t)\xi} d\xi.$$

The integrand in the second integral above is increasing in  $\xi$ . Hence the integral is bounded by the value of the integrand at  $\xi = \ln N$  times  $\ln N$  (i.e the length of the interval of integration). Thus

$$\sum_{k=3}^N \frac{(\ln k)^j}{k^t} < \frac{j^j e^{-j}}{t^j} + (\ln N)^{j+1} N^{1-t}. \quad (3.33)$$

Using (3.33) in (3.32) we get

$$Q_N^1(\alpha; j) < \int_0^1 \frac{1}{t^2} [c_j + (\ln N)^{j+1} N^{1-t} t^j] \left[ \prod_{k=3}^N \left( 1 - \frac{1}{k^t} \right) \right] dt, \quad (3.34)$$

where  $c_j := j^j e^{-j}$ . Hence,

$$Q_N^1(\alpha; j) < c_j R_N(\alpha; j) + Z_N(\alpha; j), \quad (3.35)$$

where

$$\begin{aligned} R_N(\alpha; j) &:= \int_0^1 \frac{1}{t^2} \left[ \prod_{k=3}^N \left( 1 - \frac{1}{k^t} \right) \right] dt \\ &= \int_0^1 \frac{(1 - e^{-t \ln 3})(1 - e^{-t \ln 4})}{t^2} \left[ \prod_{k=5}^N \left( 1 - \frac{1}{k^t} \right) \right] dt, \end{aligned} \quad (3.36)$$

and

$$\begin{aligned} Z_N(\alpha; j) &:= (\ln N)^{j+1} \int_0^1 N^{1-t} t^{j-2} \left[ \prod_{k=3}^N \left( 1 - \frac{1}{k^t} \right) \right] dt \\ &\leq (\ln N)^{j+1} \int_0^1 N^{1-t} \left[ \prod_{k=3}^N \left( 1 - \frac{1}{k^t} \right) \right] dt. \end{aligned} \quad (3.37)$$

Since  $R_N(\alpha; j) > 0$ , from (3.36) it is clear (by bounded convergence) that

$$\lim_N R_N(\alpha; j) = 0. \quad (3.38)$$

To treat  $Z_N(\alpha; j)$  we first use in (3.37) the estimate  $\ln(1-x) < -x$  for  $0 < x < 1$ , in order to deduce that

$$\prod_{k=3}^N \left( 1 - \frac{1}{k^t} \right) = \exp \left( \sum_{k=3}^N \ln \left( 1 - \frac{1}{k^t} \right) \right) < \exp \left( - \sum_{k=3}^N \frac{1}{k^t} \right). \quad (3.39)$$

Now,

$$\sum_{k=3}^N \frac{1}{k^t} > \int_3^N \frac{d\kappa}{\kappa^t} = \frac{N^{1-t} - 3^{1-t}}{1-t}$$

and hence (3.39) yields

$$\prod_{k=3}^N \left(1 - \frac{1}{k^t}\right) < \exp \left( -\frac{N^{1-t} - 3^{1-t}}{1-t} \right). \quad (3.40)$$

Under (3.40), formula (3.37) implies

$$Z_N(\alpha; j) < (\ln N)^{j+1} \int_0^1 N^{1-t} \exp \left( -\frac{N^{1-t} - 3^{1-t}}{1-t} \right) dt.$$

or (by substituting  $s = 1 - t$ )

$$Z_N(\alpha; j) < (\ln N)^{j+1} \int_0^1 N^s \exp \left( -\frac{N^s - 3^s}{s} \right) ds. \quad (3.41)$$

Since we are interested in letting  $N \rightarrow \infty$ , we can assume that  $N \geq 4$ . Then,  $(N^s - 3^s)/s > (\ln N - \ln 3)$  for all  $s > 0$  and hence

$$\int_0^{1/2} N^s \exp \left( -\frac{N^s - 3^s}{s} \right) ds < N^{1/2} \int_0^{1/2} e^{-(\ln N - \ln 3)} ds = \frac{3}{2} N^{-1/2}. \quad (3.42)$$

Also,

$$\int_{1/2}^1 N^s \exp \left( -\frac{N^s - 3^s}{s} \right) ds < N \frac{1}{2} e^6 \exp \left( -N^{1/2} \right). \quad (3.43)$$

Therefore, since  $Z_N(\alpha; j) > 0$ , by using (3.42) and (3.43) in (3.41) we obtain that

$$\lim_N Z_N(\alpha; j) = 0. \quad (3.44)$$

Then, since  $Q_N^1(\alpha; j) > 0$  by using (3.44) and (3.38) in (3.35) it follows that

$$\lim_N Q_N^1(\alpha; j) = 0. \quad (3.45)$$

Finally, we need to analyze  $Q_N^\infty(\alpha; j)$  of (3.31) (recall that  $a_k = \ln k$ ). Since

$$\frac{1}{1 - e^{-t \ln k}} \leq \frac{3}{2} < 2 \quad \text{for all } k \geq 3, t \geq 1,$$

we have

$$\sum_{k=3}^{\infty} (\ln k)^j \frac{e^{-t \ln k}}{1 - e^{-t \ln k}} \leq 2 \sum_{k=3}^{\infty} (\ln k)^j e^{-t \ln k} = 2 \sum_{k=3}^{\infty} \frac{(\ln k)^j}{k^t}. \quad (3.46)$$

It is now easy to check that

$$\sum_{k=3}^{\infty} \frac{(\ln k)^j}{k^t} \sim \int_1^{\infty} \frac{(\ln \xi)^j}{\xi^t} d\xi = \frac{\Gamma(j+1)}{(t-1)^{j+1}} = \frac{j!}{(t-1)^{j+1}} \quad \text{as } t \rightarrow 1^+. \quad (3.47)$$

Hence, from (3.46) and (3.47) it follows that there is a constant  $M > 0$  such that

$$\sum_{k=3}^{\infty} (\ln k)^j \frac{e^{-t \ln k}}{1 - e^{-t \ln k}} \leq \frac{M}{(t-1)^{j+1}} \quad \text{for all } t > 1. \quad (3.48)$$

In addition, for  $t \rightarrow \infty$  we have estimate (3.11), namely

$$\sum_{k=3}^{\infty} (\ln k)^j \frac{e^{-t \ln k}}{1 - e^{-t \ln k}} = O\left(\frac{1}{3^t}\right). \quad (3.49)$$

Estimates (3.48) and (3.49) imply that, for  $N \geq j + 4$ ,

$$\int_1^{\infty} \left[ \sum_{k=3}^{\infty} (\ln k)^j \frac{e^{-t \ln k}}{1 - e^{-t \ln k}} t^{j-1} \right] \left[ \prod_{k=3}^N (1 - e^{-t \ln k}) \right] dt < \infty. \quad (3.50)$$

We, therefore, have a case similar to the one of Proposition 3: For  $N \geq j + 4$  it holds that

$$\begin{aligned} & \left[ \sum_{k=3}^N (\ln k)^j \frac{e^{-t \ln k}}{1 - e^{-t \ln k}} t^{j-1} \right] \prod_{k=3}^N (1 - e^{-t \ln k}) \\ & < \left[ \sum_{k=3}^{\infty} (\ln k)^j \frac{e^{-t \ln k}}{1 - e^{-t \ln k}} t^{j-1} \right] \prod_{k=3}^{j+4} (1 - e^{-t \ln k}) \end{aligned} \quad (3.51)$$

and, consequently, under (3.50), dominated convergence implies

$$\lim_N Q_N^{\infty}(\alpha; j) = \int_1^{\infty} \left[ \sum_{k=3}^{\infty} (\ln k)^j \frac{e^{-t \ln k}}{1 - e^{-t \ln k}} t^{j-1} \right] \left[ \prod_{k=3}^{\infty} (1 - e^{-t \ln k}) \right] dt < \infty. \quad (3.52)$$

Finally, combining (3.45) and (3.52) with (3.29) we obtain

$$\lim_N E[U_j^N] = I(\alpha; j) = \frac{1}{(j-1)!} \int_1^{\infty} \left[ \sum_{k=3}^{\infty} \frac{(\ln k)^j}{k^t} \frac{t^{j-1}}{1 - e^{-t \ln k}} \right] \left[ \prod_{k=3}^{\infty} \left(1 - \frac{1}{k^t}\right) \right] dt < \infty.$$

Notice that the quantity  $I(\alpha; j)$  of the above formula is the same as that of (3.14) (with  $\alpha = \{\ln k\}_{k=3}^{\infty}$ ). This completes the discussion of Example 6.

**Acknowledgment.** The second author (VGP) is grateful to the Boeing Center for Technology, Information and Manufacturing (BCTIM) of the Olin School of Business, Washington University in St. Louis, and its director, professor P. Kouvelis, for their hospitality. This paper was written during the author's visit at BCTIM.

## References

- [1] I. Adler, S. Oren and S. Ross, The coupon collector’s problem revisited, *J. Appl. Prob.* **40** (2003) 513–518.
- [2] L.E. Baum and P. Billingsley, Asymptotic distributions for the Coupon Collector’s problem, *Annals of Mathematical Statistics* **36** (1965) 1835–1839.
- [3] J. D. Boisserranger, D. Gardy, and Y. Ponty, The weighted words collector, *DMTCS Proceedings, 23rd Intern. Meeting on Probabilistic, Combinatorial, and Asymptotic Methods for the Analysis of Algorithms (AofA ’12)*, Canada, Montreal, (2012).
- [4] A. Boneh and M. Hofri, The Coupon Collector Problem Revisited—a Survey of Engineering Problems and Computational Methods, *Comm. Statist. Stochastic Models* **13** (no. 1) (1997) 39–66.
- [5] S. Boneh and V.G. Papanicolaou, General Asymptotic Estimates for the Coupon Collector Problem, *Journal of Computational and Applied Mathematics* **67** (2) (Mar. 1996) 277–289.
- [6] G. Boros and V.H. Moll, *Irresistible Integrals: Symbolics, Analysis and Experiments in the Evaluation of Integrals*, Cambridge University Press, Edinburgh, 2004.
- [7] R.K. Brayton, On the asymptotic behavior of the number of trials necessary to complete a set with random selection, *Journal of Mathematical Analysis and Applications* **7** (1963) 31–61.
- [8] P. Diaconis and S. Holmes, A Bayesian peek into Feller volume I, *Sankhyā*, Special issue in memory of D. Basu, **64** Ser. A (3, part 2) (2002) 820–841.
- [9] A.V. Dumas and V.G. Papanicolaou, The Coupon Collector’s Problem Revisited: Asymptotics of the Variance, *Adv. Appl. Prob.* **44** (1) (2012) 166–195.
- [10] A.V. Dumas and V.G. Papanicolaou, Asymptotics of the rising moments for the coupon collector’s problem, *Electron. J. Probab.* **Vol. 18** (Article no. 41) (2012) 1–15 (DOI: 10.1214/EJP.v18-1746).
- [11] R. Durrett, *Probability: Theory and Examples*, Third Edition, Duxbury Advanced Series, Brooks/Cole—Thomson Learning. Belmont, CA, USA, 2005.
- [12] P. Erdős and A. Rényi, On a classical problem of probability theory, *Magyar. Tud. Akad. Mat. Kutató Int. Közl.*, **6** (1961), 215–220.



- [13] W. Feller, *An Introduction to Probability Theory and Its Applications*, Vol. I & II, John Wiley & Sons, Inc., New York, 1966.
- [14] D. Foata, H. Guo-Niu and B. Lass, Les nombres hyperharmonique et la fratrie du collectionneur de vignettes, *Sem. Lothar Combinatoire* **47** B47a (2001) (electronic).
- [15] D. Foata and D. Zeilberger, The Collector's Brotherhood Problem using the Newman–Shepp symbolic method, *Algebra univers.* **49** (2003) 387–395.
- [16] H.J. Goodwin, On cartophily and motor cars, *Math. Gazette* **33** (305) (1949) 169–171.
- [17] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Fourth Edition, Oxford University Press, Glasgow, 1960.
- [18] L. Holst, On Birthday, Collectors', Occupancy and other classical Urn problems, *International Statistical Review* **54** (1986) 15–27.
- [19] S. Janson, Limit theorems for some sequential occupancy problems, *Journal of Applied Probability* **20** (1983) 545–553.
- [20] H.M. Mahmoud, *Pólya urn models*, CRC Press, New York, 2008.
- [21] F.G. Maunsell, A problem in cartophily, *Math. Gazette* **22** (251) (1938) 328–331.
- [22] P. Neal, The Generalized Coupon Collector Problem, *J. Appl. Prob.* **45** (2008) 621–629.
- [23] D.J. Newman and L. Shepp, The Double Dixie Cup problem, *American Mathematical Monthly* **67** (1960) 58–61.
- [24] N. Pintacuda, Coupons collectors via the martingales, *Bolletino Un. Mat. It.* **17A** (5) (1980) 174 – 177.
- [25] S. Ross, *A First Course in Probability*, Eighth Edition, Pearson Prentice Hall, Pearson Education, Inc., Upper Saddle River, NJ, 2010.
- [26] S. Ross, *Introduction to Probability Models*, Tenth Edition, Elsevier Inc., Burlington, MA, 2010.
- [27] W. Rudin, *Real and Complex Analysis*, Third Edition, McGraw-Hill, New York, 1987.
- [28] The Dixie Cup Company History:  
<http://academicmuseum.lafayette.edu/special/dixie/company.html>